Energy functions for dissipativity-based balancing of discrete-time nonlinear systems

Lopezlena, Ricardo; Scherpen, Jacquelien M.A.

Published in:
Mathematics of Control, Signals, and Systems (MCSS)

DOI:
10.1007/s00498-006-0007-z

Math. Control Signs Syst. (2006) 18:345–368
DOI 10.1007/s00498-006-0007-z

ORIGINAL ARTICLE

Energy functions for dissipativity-based balancing of discrete-time nonlinear systems

Ricardo Lopezlena · Jacquelien M. A. Scherpen

Received: 13 December 2003 / Revised: 15 June 2005 /
Published online: 5 August 2006

Abstract Most of the energy functions used in nonlinear balancing theory can be expressed as storage functions in the framework of dissipativity theory. By defining a framework of discrete-time dissipative systems, this paper presents existence conditions for their discrete-time energy functions along with algorithms to find them based on dynamic optimization problems. Furthermore, the important case of the nonlinear discrete-time versions of the controllability and observability functions, its properties and algorithms to find them are presented. These algorithms are illustrated with linear and nonlinear examples.

Keywords Nonlinear systems · Dissipative systems · Discrete-time systems · Controllability · Observability

1 Introduction

The study of systematic tools for model reduction of dynamic systems has been an early topic of interest in the systems and control fields. Model approximation based on the Hankel norm and the balancing method [4,18] have shown to be useful tools for model reduction for linear systems. Furthermore, singular values-based balancing, LQG balancing and $\mathcal{H}_\infty$ balancing are nowadays important tools for linear model reduction. With the use of the behavioral approach, in [30] a general balancing framework for model approximation and

R. Lopezlena (✉) · J. M. A. Scherpen
Delft Center for Systems and Control, Delft University of Technology,
Mekelweg 2, 2628 CD Delft, The Netherlands
e-mail: R.Lopezlena@dcsc.TUDelft.nl

J. M. A. Scherpen
e-mail: J.M.A.Scherpen@dcsc.TUDelft.nl

reduction is provided, which has Lyapunov, LQG, and $\mathcal{H}_\infty$ as special cases being valid for linear unstable systems [30]. These developments provide interesting paradigms for nonlinear generalizations.

In continuous-time nonlinear systems, there has been important progress on the nonlinear extensions of balancing methods, mainly based on the controllability and observability functions and their corresponding singular-values [3,23, 24], but also with other energy functions [5] or for particular problems, namely LQG [26], $\mathcal{H}_\infty$ [25] or for port-Hamiltonian Systems [15]. The use of the theory of dissipative systems offers to a certain extent a generalized approach in order to deal with the variety of energy functions used for nonlinear balancing of continuous systems [15]. The explicit use of dissipativity theory for the balancing of linear systems was first presented in [29].

Most of the efforts have been devoted to continuous-time systems. The prototypical case is precisely nonlinear balancing based on the controllability and observability functions. Roughly speaking, in the procedure presented in [23], a Hamilton–Jacobi equation and a Lyapunov-like partial differential equation have to be solved in order to determine the energy functions. Then with the use of a nonlinear coordinate transformation, the system is represented in a balanced form. After truncation of the less important dynamic subsystem and application of an inverse transformation a reduced system is obtained. The mathematical complexity in solving the partial differential equations associated to the energy functions has stimulated the search for alternative methods [20].

In this paper some aspects of the discrete-time framework for dissipativity theory for balancing nonlinear systems are introduced. Such framework relies on *storage functions*, in particular the *required supply* and the *available storage*, instead of the controllability and observability functions. Therefore, in order to find such storage functions, optimization algorithms are proposed. Furthermore, the energy functions for stable nonlinear discrete-time systems are discussed as important particular cases, extending the continuous-time theory presented in [23,24]. Since the determination of such storage functions are a fundamental condition for nonlinear balancing and model reduction, in this paper attention is given to computer implementation of the theory. Furthermore, with applications in mind, some preliminary connections with continuous systems that are time-discretized are presented. This approach does not assume any linearization procedure at all, contrasting with [28]. In the preliminary work [14] it was shown that once the energy functions are found, some procedures for the continuous-time balancing presented in [3,24] can be followed in order to find a reduced system.

The paper is organized as follows. After fixing the notation used, in Sect. 2, relationships with time-discretized systems are presented. These concepts allow us to represent in a different form an optimal control problem that appears in the following Sect. 3, where several concepts of the discrete-time dissipativity theory are discussed. One important case of the storage functions presented in Sect. 3 are the discrete-time energy functions. In Sect. 4, the observability and controllability functions and their properties are discussed and algorithms to find them are presented. In order to illustrate these methods, linear and nonlinear examples are shown and briefly discussed. Finally, some conclusions are presented.

Notation: Despite the efforts of several experts [8,16,17], notation for nonlinear discrete-time systems is not standard. The notation in this paper tends to follow [12,13]. The set of nonnegative and nonpositive integers are denoted as $\mathbb{Z}^+ \stackrel{\text{def}}{=} \{0, 1, 2, \ldots\}$ and $\mathbb{Z}^- \stackrel{\text{def}}{=} \{0, -1, -2, \ldots\}$, respectively. Time is denoted by $t \in \mathbb{R}^+$ and while $T \in \mathbb{R}^+$ denotes the endpoint of a certain finite period of time, $\mathcal{T} \in \mathbb{R}^+$ denotes the sampling time, $t = k\mathcal{T}$, $k \in \mathbb{Z}^+$. Discrete-time vector variables are denoted for instance as x_k or $x(k)$. A continuous-time function is expressed as $f(t)$, which is expressed after discretization by the discrete-time function $F(k)$. Where convenient, for clarity of exposition a function of several variables $F(x_k, u_k)$ may be denoted simply as F_k. Given a function F_k its inverse function (map) is denoted as F_k^{-1}. An optimal input variable at time k is denoted as $v_k^\star$. Finally, the solution of the system $x_{k+1} = F(x_k, u_k)$ at the interval $[k, k+m]$ with initial condition $x(k) = x_k$ and input $u_k \in \ell_2(0, \infty)$ is denoted by $x_{k+m} = \phi(k+m, k, x_k, u)$.

2 Some relationships between continuous and discrete-time systems

In this section we begin to study several relations of continuous and discrete-time systems (and other associated systems that can be derived from them) with the purpose of simplifying the interpretation of one optimal control problem that appears in the subsequent Sect. 3.

Consider the following continuous-time nonlinear system

$$\dot{x}(t) = f(x(t), u(t)), \tag{1}$$
$$y(t) = h(x(t)), \tag{2}$$

where $x \in \mathbb{R}^n$ are local coordinates for a C^∞ state space manifold $\mathcal{X}$, f and h are C^∞, $u \in \mathcal{U} \subset \mathbb{R}^p$ and $y \in \mathcal{Y} \subset \mathbb{R}^q$. Assume throughout that u, x and y are locally square integrable. On the other side, consider the following discrete-time nonlinear system,

$$x_{k+1} = F(x_k, u_k), \tag{3}$$
$$y_k = h(x_k), \qquad k \in \mathbb{Z} \tag{4}$$

where $u_k = (u_1, \ldots, u_p)_k \in \mathcal{U} \subset \mathbb{R}^p$, $y_k = (y_1, \ldots, y_q)_k \in \mathcal{Y} \subset \mathbb{R}^q$ and $x_k = (x_1, ..., x_n)_k \in \mathbb{R}^n$ are local coordinates for the smooth state space manifold $\mathcal{X}$. Moreover, F and h are class C^∞ in a neighborhood $D \subset \mathbb{R}^n$ around an equilibrium point in $x = 0$ such that $F(0, 0) = 0$ and $h(0) = 0$.

In discrete-time systems, evolution in reversed-time implies the invertibility of the map $F(\cdot, u)$, which is only possible under certain conditions, discussed in [2] and [8]. Roughly speaking, any discrete-time nonlinear system that is *causal* can be described by a generically reversible dynamics [2], and when sampling or discretizing a system in the form (1) to obtain (3) the resulting dynamics is reversible, meaning that the Jacobian matrix $[\partial F/\partial x, \partial F/\partial u]$ is generically nonsingular for all values of x and u, [2,8], and the system is said to be generically submersive [7].

Some problems in optimal control can be simplified when the time direction of evolution is reversed, like in the cases discussed in the following Sects. 2 and 3 where we will assume that $F(\cdot, u)$ is a diffeomorphism. If it is the case that the system (3) results from discretization of the continuous-time system (1) then this latter assumption is satisfied automatically, [2,8]. There are several publications regarding the procedures to obtain a discrete-time equivalent of the state Eq. (1), e.g., [9,10,16]. In particular, the method known as the *Taylor–Lie series* discretization [9], yields a system in the form of (3) for $k \in \mathbb{Z}^+$. The Taylor–Lie series discrete-time equivalent to (1) is given by

$$x_{k+1} = x_k + \sum_{i=1}^{\infty} \frac{\mathcal{T}^i}{i!} \left(\frac{\mathrm{d}^i x}{\mathrm{d}t^i} \right)_k, \tag{5}$$

where $\mathcal{T}$ is the sampling-time and

$$\frac{\mathrm{d}^i x}{\mathrm{d}t^i} = \begin{cases} f(x(t), u(t)), & \text{for } i = 1, \\ \frac{\partial^T}{\partial x} \left(\frac{\mathrm{d}^{i-1} x}{\mathrm{d}t^{i-1}} \right) f(x(t), u(t)), & \forall i > 1. \end{cases} \tag{6}$$

This discretization procedure preserves several analytical properties like equilibrium and is a generalization of the exponential matrix discretization procedure for linear systems.

When needed, such procedure will be used to establish certain relations between continuous and time-discretized systems, which allows us to count with alternative representations of certain time-reversed discrete-time nonlinear optimal control problems in the following sections. An exact discretization from Eq. (5) is assumed in the sense that the infinite series is considered without truncation during calculations.

Example 2.1 Consider the linear time invariant, stable, minimal, continuous-time system

$$\begin{aligned} \dot{x}(t) &= \mathcal{A}x(t) + \mathcal{B}u(t), \\ y(t) &= Cx(t), \end{aligned} \tag{7}$$

which by the Taylor–Lie method (or by the step-invariant discretization procedure) results in the following discrete-time system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k, \\ y_k &= Cx_k, \end{aligned} \tag{8}$$

where $u \in \mathbb{R}^p$, $y \in \mathbb{R}^q$ and $x \in \mathbb{R}^n$, $A = \mathrm{e}^{\mathcal{A}\mathcal{T}}$ and $B = (\sum_{i=1}^{\infty} \frac{(\mathcal{A})^{i-1}\mathcal{T}^i}{i!})\mathcal{B} = \int_0^{\mathcal{T}} \mathrm{e}^{\mathcal{A}s}\mathcal{B}\,\mathrm{d}s$.

In order to distinguish stringently the direction of time evolution throughout the paper, define the past interval $I_p = [-T_p, 0] \in \mathbb{R}^1$, $T_p > 0$ the future interval $I_f = [0, T_f] \in \mathbb{R}^1$, $T_f > 0$ and the total interval $I = [-T_p, T_f] \in \mathbb{R}^1$, $T_f + T_p = T_I > 0$. A forward-time evolution is denoted by $t \in I$ (a strictly increasing variable from $t_0 \in I$) and a backward-time evolution with $\tau \in I$ (a strictly decreasing variable from $\tau_0 \in I$). A sequence evolving in forward-time is denoted by $k \in \mathbb{Z}$ such that $-N_p \leq k \leq N_f$ from some initial $k_o \in \mathbb{Z}$ and a backward-time evolution by $\kappa \in \mathbb{Z}$ such that $-N_p \leq \kappa \leq N_f$ from some initial $\kappa_o \in \mathbb{Z}$; $N_p, N_f \in \mathbb{Z}^+$.

Definition 2.1 *Assume that system* (1) *evolves in $t \in [-T_p, 0]$. Define an associated system by*

$$\frac{\mathrm{d}}{\mathrm{d}t} w(t) = -f(w(t), v(t)), \quad t \in [0, T_p], \tag{9}$$

which can be obtained by performing two sequential operations on Eq. (1):
Backward-time: *By rendering Eq.* (1) *evolve in regressive time i.e., $\tau = -t$, $x(-\tau) = x(t)$, $\mathrm{d}x(-\tau)/\mathrm{d}\tau = -\mathrm{d}x(t)/\mathrm{d}t$ and $u(-\tau) = u(t)$.*
Flip-time: *By defining $w(t) \stackrel{\text{def}}{=} x(-\tau)$ and $v(t) \stackrel{\text{def}}{=} u(-\tau)$.*

In the same form an associated discrete-time system can be defined as follows:

Definition 2.2 *Assume that departing from $k = -N_p$, the system* (3) *evolves in $-N_p \leq k \leq 0$. Define an associated system by*

$$w_{k+1} = F^{-1}(w_k, v_{k+1}), \quad 0 \leq k \leq N_p \tag{10}$$

which departs from $k = 0$ and can be obtained by applying two sequential operations on Eq. (3):
Backward-time: *Inverting the map in Eq.* (3) *replacing the time index by κ and, evolving, from* 0*, with the decreasing sequence $\kappa \in \mathbb{Z}^-$.*
Flip-time: *By replacing the time index as $k := -\kappa$, $0 \leq k \leq N_p$ and (mirrored with respect to* 0*) defining $w_k \stackrel{\text{def}}{=} x_\kappa$ (thus $w_{k-1} := x_{\kappa+1}$) and $v_k \stackrel{\text{def}}{=} u_\kappa$.*

The steps detailed in the continuous and discrete-time systems under Definitions 2.1 and 2.2, respectively, involve the realization of a sequence of associated systems, which are related to each other by a discretization procedure, as shown in the commutative diagram of Fig. 1.

Remark 2.1 Commutativity of the diagram presented in Fig. 1 depends on the fact that the discrete-time systems represented there, depart from continuous-time systems and therefore the map $F(\cdot, u)$ is invertible [8]. For practical applications several ways to obtain the state of the time-discretized system can be used, including some algorithms of numerical integration.

$$
\begin{tikzcd}
\dot{x}(t) = f(x(t), u(t)) \arrow[r, "\text{Discretization}"] \arrow[d, "\tau=-t"'] & x_{k+1} = F(x_k, u_k) \arrow[d, "\kappa"'] \\
\dot{x}(-\tau) = -f(x(-\tau), u(-\tau)) \arrow[r, "\text{Discretization}"] \arrow[d, "\text{flip}"] & x_\kappa = F^{-1}(x_{\kappa+1}, u_\kappa) \arrow[d, "\text{flip}"', "k=-\kappa"] \\
\dot{w}(t) = -f(w(t), v(t)) \arrow[r, "\text{Discretization}"] & w_{k+1} = F^{-1}(w_k, v_{k+1})
\end{tikzcd}
$$

Fig. 1 Commutative diagram

Remark 2.2 In particular, the associated system (10) can be found alternatively by time-discretizing (1) and performing the operations in Definition 2.2, or alternatively by direct time-discretization of system (9), resulting in the following system

$$w_{k+1} = w_k + \sum_{i=1}^{\infty} \frac{\mathcal{T}^i}{i!} \left(\frac{\mathrm{d}^i w}{\mathrm{d}t^i} \right)_k \tag{11}$$

with

$$\frac{\mathrm{d}^i w}{\mathrm{d}t^i} = \begin{cases} -f(w(t), v(t)), & \text{for } i = 1 \\ (-1)^i \frac{\partial^T}{\partial w} \left(\frac{\mathrm{d}^{i-1} w}{\mathrm{d}t^{i-1}} \right) f(w(t), v(t)), & \forall i > 1, \end{cases} \tag{12}$$

and defining $v(k+1)$ as the time discretization of $v(t)$, which circumvents the inversion of the map $F(\cdot, u)$.

In the next section an optimal control problem will have a simpler solution due to the tools we have presented in this section. This section concludes with the linear system example to illustrate these concepts.

Example 2.2 Consider the system (8) in Example 2.1 which by Def. 2.2 it has an associated system given by

$$w_{k+1} = A^{-1} w_k - A^{-1} B v_{k+1}, \tag{13}$$

since $A^{-1}B = \mathrm{e}^{-\mathcal{A}\mathcal{T}} \int_0^{\mathcal{T}} \mathrm{e}^{\mathcal{A}s} \mathcal{B}\, ds$ and by defining $\xi = \mathcal{T} - s$, $\mathrm{d}\xi = -\mathrm{d}s$, it can be alternatively written in terms of $(\mathcal{A}, \mathcal{B})$ as

$$w_{k+1} = \mathrm{e}^{-\mathcal{A}\mathcal{T}} w_k - \int_0^{\mathcal{T}} \mathrm{e}^{-\mathcal{A}\xi} \mathcal{B}\, \mathrm{d}\xi \; v_{k+1}. \tag{14}$$

Consider now the use of the commutative diagram in Fig.1 along with Remark 2.2. Since the use of Definition 2.1 in system (7) provides us with the continuous associated system $\dot{w}(t) = -\mathcal{A}w(t) - \mathcal{B}v(t)$, $t \geq 0$, the discretization of this system yields straightforwardly a system in the form of (8) with $A = \mathrm{e}^{-\mathcal{A}\mathcal{T}}$ and $B = -\int_0^{\mathcal{T}} \mathrm{e}^{-\mathcal{A}\xi}\mathcal{B}\,\mathrm{d}\xi$, which is Eq. (14).

3 Discrete-time dissipativity theory and storage functions

In [15] a framework based on dissipativity theory for balancing and nonlinear model reduction of continuous systems was presented. In this section, such framework is presented in its *discrete-time* form along with some results whose proofs (direct equivalence with the continuous-time, e.g., [27,31]) provide the way to reinterpret some optimal control problems as dynamic optimization problems. Some concepts of the discrete-time theory of dissipative systems have been developed elsewhere [6,12].

The system (3) is said to be dissipative with *supply rate* $r(y_k, u_k)$, $r : \mathcal{Y} \times \mathcal{U} \to \mathbb{R}$, if there exists a nonnegative function $S : \mathbb{R}^n \to \mathbb{R}$, $S(0) = 0$ called *storage function* such that for all $u_k \in \mathcal{U}$ and all $k \in \mathbb{Z}$ [12],

$$S(x_{k+1}, r) - S(x_k, r) \leq r(y_k, u_k), \tag{15}$$

which for all k, $m \in \mathbb{Z}$, with $m \geq 0$ is equivalent to

$$S(x_{k+m}, r) - S(x_k, r) \leq \sum_{i=k}^{k+m-1} r(y_i, u_i). \tag{16}$$

This latter relation is named the *discrete-time dissipation inequality* [12,31].

Theorem 3.1 *The system* (3)–(4) *is dissipative with supply rate* $r_a(y_k, u_k)$ *if and only if the function called available storage,* $S_a : \mathcal{X} \to \mathbb{R}^+$*, defined as*

$$\begin{aligned} S_a(x^0, r_a) &= \sup_{\substack{u(\cdot)\in\mathcal{U} \\ x(0)=x^0,\, N_f\in\mathbb{Z}^+}} -\sum_{i=0}^{N_f} r_a(y_i, u_i) \\ &= -\inf_{\substack{u(\cdot)\in\mathcal{U} \\ x(0)=x^0,\, N_f\in\mathbb{Z}^+}} \sum_{i=0}^{N_f} r_a(y_i, u_i), \quad i \in \mathbb{Z}^+. \end{aligned} \tag{17}$$

is finite for all $x^0 \in \mathcal{X}$*. In such case* $S_a(x^0, r_a)$ *is a storage function such that* $S_a(x^0, r_a) \leq S(x^0, r)$ *for all* $x^0 \in \mathcal{X}$*, for any other storage function* $S(x^0, r)$*.*

The proof of Theorem 3.1 can be found in the Appendix.

Lemma 3.1 *Assume that there exists the optimal sequence of inputs* $\{u_i^\star | i = 0, 1, \ldots, N_f - 1\}$ *that fulfills* (17). *Then*

$$S_a(x^0, r_a) = -\sum_{i=0}^{N_f} r_a(y_i, u_i^\star), \tag{18}$$

and moreover, it can be found from the limit $S_a(x^0, r_a) = \lim_{k \to N_f} \mathcal{S}_a(x_{k+1}, r_a)$ *where* $\mathcal{S}_a(x_{k+1}, r_a)$ *is the solution of the following recurrence equation*

$$\mathcal{S}_a(x_{k+1}, r_a) = \mathcal{S}_a(x_k, r_a) - r_a(y_k, u_k^\star) \tag{19}$$

with initial condition $\mathcal{S}_a(x^0, r_a) = 0$.

Proof The simple result of Lemma 3.1 can be shown by solving the iterative Eq. (19) with the initial condition provided. □

Consider the second representation in Eq. (17). With the assumptions of Theorem 3.1, the optimal control problem (17) can be interpreted as a dynamic optimization algorithm as follows. Let $\epsilon, N_f \in \mathbb{Z}^+$ be such that $\|x_{N_f}\| \leq \epsilon$ for ϵ small enough. Assume that N_f is known and that the (closed) set of admissible inputs $\{u \mid u \in \mathcal{U}\}$ (each $u \in \mathcal{U}$ being locally square integrable) satisfy the usual regularity conditions of being convex with nonempty interior. Then $S_a(x^0, r_a)$ as given by (17) restricted to system (3)–(4) with boundary conditions $x_{N_f} = 0$ and $x^0 = x(0)$ can be posed as the solution of the following dynamic optimization problem

$$\min_{\{u_i | i=0,\ldots,N_f-1\}} \mathcal{S}_a(x_{N_f}, r_a), \tag{20}$$

with equality constraints

$$\left\{ \begin{aligned} x_{i+1} &= F(x_i, u_i), \\ \mathcal{S}_a(x_{i+1}, r_a) &= \mathcal{S}_a(x_i, r_a) + r_a(h(x_i), u_i), \\ x_{N_f} &= 0, \\ x_0 &= x^0, \end{aligned} \right. \tag{21}$$

with initial inputs $\{u^0 \mid u_j^0 \in \mathcal{U},\ j = 0, 1, \ldots, N_f - 1\}$ and with $\mathcal{S}_a(x_0, r_a) = r_a(h(x^0), u_0^\star)$. With the determination of $u^\star$, then $S_a(x_k, r_a) = -\mathcal{S}_a(x_k, r_a)$.

Theorem 3.2 *Assume that* (3) *is reachable from* $x^* \in \mathcal{X}$, *then the required supply,* $S_r : \mathcal{X} \to \mathbb{R}^+$, *defined as*

$$S_r(x^0, r_r) = \inf_{\substack{u(\cdot) \in \mathcal{U} \\ x(0)=x^0,\ N_p \in \mathbb{Z}^+}} \sum_{i=-N_p}^{0} r_r(y_i, u_i), \qquad i \in \mathbb{Z}^-, \tag{22}$$

satisfies the dissipation inequality (16). *Then* (3) *is dissipative if and only if* $S_r(x^0, r_r)$ *is finite for all* $x \in \mathcal{X}$.

The proof of Theorem 3.2 can be found in the Appendix.

The definition of the associated system (10) provides a way to express the same optimal control problem defining $S_r(x, r_r)$ but in forward time, convenient for the subsequent developments. During the rest of the paper from Theorem 3.2 it will be assumed that $x^*_{-N_p} = 0$.

Remark 3.1 Taking in consideration the system (10) from Definition. 2.2, then S_r from Eq. (22), may be alternatively expressed as

$$S_r(w^0, r_r) = \inf_{\substack{v(\cdot)\in\mathcal{U} \\ w^0=x(0),\, N_p\in\mathbb{Z}^+}} \sum_{i=0}^{N_p} r_r(y_i, v_i), \tag{23}$$

for w_k and v_k as in Definition 2.2.

Remark 3.2 v_0 does not influence the new state w_1 in (10), where it results $w_1 = F^{-1}(w_0, v_1)$. Therefore the value of v_0 which minimizes (23) is $v_0^\star = 0$ and thus $u_0^\star = 0$.

Lemma 3.2 *Assume the existence of the optimal sequence* $v^\star = \{v_i^\star | i = 0, 1, \ldots N_p\}$ *such that it satisfies* (23) *and consider the following recursive equation*

$$\mathcal{S}_r(w_{i+1}, r_r) = \mathcal{S}_r(w_i, r_r) + r_r(y_i, v_i), \tag{24}$$

for $i = 0, 1, 2, \ldots$ *and initial condition* $\mathcal{S}_r(w^0, r_r) = r_r(y_0, v_0^\star)$. *Then* $S_r(w^0, r_r)$ *can be found from the iterative solution of Eq.* (24).

Proof Express (23) as,

$$S_r(w^0, r_r) = r_r(y_0, v_0^\star) + \sum_{i=0}^{N_p} r_r(y_i, v_i^\star), \tag{25}$$

which may be written as a recurrence equation with the initial condition $\mathcal{S}_r(w^0, r_r) = r_r(y_0, v_0^\star)$ as consequence of Remark 3.2. By solving iteratively (24), $S_r(w^0, r_r)$ can be found as i tends to N_p. □

Assuming that the conditions of Theorem 3.2 hold, then the approximate solution of the optimal control problem (22) can be found by a reinterpretation of the problem as follows. Let $\epsilon, N_p \in \mathbb{Z}^+$ be such that $\|w_{N_p}\| \leq \epsilon$ for ϵ small enough. Assume that N_p is known and assume the following regularity condition is satisfied: the (closed) set of admissible inputs $\{v \mid v \in \mathcal{V}\}$ is convex with nonempty interior. Then $S_r(w_0, r_r)$ as given by (22) restricted to (10) with boundary conditions $w_{N_p} = 0$ and $w^0 = w_k$ can be posed as the solution of the following dynamic optimization problem

$$\min_{\{v_i | i=1,\ldots,N_p\}} \mathcal{S}_r(w_{N_p}, r_r), \tag{26}$$

with equality constraints

$$\begin{cases} w_{i+1} = F^{-1}(w_i, v_{i+1}), \\ \mathcal{S}_r(w_{i+1}, r_r) = \mathcal{S}_r(w_i, r_r) + r_r(h(w_i), v_i), \\ w_{N_p+1} = 0, \\ w^0 = w_0, \end{cases} \tag{27}$$

with initial inputs $\{v^0 \mid v_j^0 \in \mathcal{V},\ j = 1, \ldots, N_p\}$ and with $\mathcal{S}_r(w^0, r_r) = r_r(h(w^0), 0)$, determining $v_i^\star$.

4 Discrete-time controllability and observability functions

In this section we restrict ourselves to the important case in the context of dissipative systems, with the required supply S_r with supply rate $r_r = \frac{1}{2}u_k^T u_k$ and the available storage S_a with supply rate $r_a = -\frac{1}{2}y_k^T y_k$.

Definition 4.1 *The controllability and observability functions of the system* (3) *are defined respectively as,*

$$L_c(x^0) = \inf_{\substack{u \in \ell_2(-\infty,0), \\ x(-\infty)=0,\ x^0=x_0}} \frac{1}{2} \sum_{k=-\infty}^{0} \| u_k \|^2, \quad k \in \mathbb{Z}^- \tag{28}$$

$$L_o(x^0) = \frac{1}{2} \sum_{k=0}^{\infty} \| y_k \|^2,\ x^0 = x_0,\ u_k = 0, \quad k \in \mathbb{Z}^+. \tag{29}$$

The energy functions (28) and (29) so defined are the discrete-time equivalents of the continuous versions presented in [23]. When these functions are defined for linear systems like Eq. (8) some known functions result in terms of Gramians [22].

Corollary 4.1 *Consider the system* (8). *Then L_c and L_o, as defined in Eqs.* (28) *and* (29)*, are given by,*

$$L_c(x^0) = \frac{1}{2} x_0^T P^{-1} x_0, \tag{30}$$

$$L_o(x^0) = \frac{1}{2} x_0^T Q x_0, \tag{31}$$

with Gramians $P = \sum_{k=0}^{\infty} A^k B B^T A^{kT}$ and $Q = \sum_{k=0}^{\infty} A^{kT} C^T C A^k$.

The proof of this result is presented in the examples throughout this section.

Remark 4.1 Assuming that system (3) is dissipative for a supply rate $r_a = -\frac{1}{2}y_k^T y_k$ and $N_f \to \infty$, the discrete-time version of the *new* or *generalized* observability energy function defined by Gray and Mesko [5]

$$\hat{L}_o(x^0) = \sup_{\substack{u\in\ell_2(0,\infty),\ \|u\|_{\ell_2}\leq\alpha \\ x_0=x^0,\ x(\infty)=0}} \frac{1}{2}\sum_{k=0}^{\infty} \| y_k\|^2, \quad \alpha \geq 0,\ k \in \mathbb{Z}^+,$$

can be considered a particular case of (17) and furthermore can be posed as a dynamic optimization algorithm in the form of (20) and (21).

During the rest of this section we restrict ourselves to the functions (28) and (29) resulting from system (3). From now on assume system (3) to be locally asymptotically stable at $x^0 = x_0$ for $u_k = 0$, in a neighborhood $D \subset \mathbb{R}^n$.

4.1 Observability function

Since the definition of L_o, Eq. (29) does not define an optimal control problem, in this subsection a recursive procedure to find the observability function is provided along with some properties. Also a Lyapunov-like difference equation analog to that found in [24], is presented.

Lemma 4.1 *Consider the following recursive equation*

$$\mathcal{L}_o(x_{i+1}) = \mathcal{L}_o(x_i) + \frac{1}{2}h^T[F(x_i,0)]h[F(x_i,0)], \tag{32}$$

for $i = 0, 1, 2\dots$ *and* $\mathcal{L}_o(x^0) = \frac{1}{2}h^T(x^0)h(x^0)$ *as initial condition. Then* $L_o(x^0)$ *can be found from the solution of (32) as follows*

$$L_o(x^0) = \lim_{i\to\infty} \mathcal{L}_o(x_i). \tag{33}$$

Proof By Eq. (3) and by definition of L_o in Eq. (29), one obtains

$$L_o(x^0) = \frac{1}{2}\sum_{i=0}^{\infty} h^T(x_i)h(x_i) = \frac{1}{2}h^T(x^0)h(x^0) + \frac{1}{2}\sum_{i=0}^{\infty} h^T[F(x_i,0)]h[F(x_i,0)], \tag{34}$$

for $i \in \mathbb{Z}^+$. Noting that

$$\mathcal{L}_o(x_{N_f+1}) = \mathcal{L}_o(x^0) + \frac{1}{2}\sum_{i=0}^{N_f} h^T[F(x_i,0)]h[F(x_i,0)], \quad N_f \in \mathbb{Z}^+$$

the result is obtained from Eq. (33) when $N_f \to \infty$. □

Example 4.1 An alternative proof of Corollary 4.1 to find L_o follows. Use recurrent Eq. (32) for the system (8) resulting in the following difference equation

$$\mathcal{L}_o(x_{i+1}) = \mathcal{L}_o(x_i) + \frac{1}{2}x_i^T A^T C^T C A x_i, \tag{35}$$

with initial condition $\mathcal{L}_o(x_0) = \frac{1}{2}x_0^T C^T C x_0$. Then the solution of (35) yields

$$L_o(x^0) = \lim_{i\to\infty} \mathcal{L}_o(x_i) = \sum_{k=0}^{\infty} x_0^T A^{kT} C^T C A^k x_0,$$

which is Eq. (31).

Definition 4.2 *System* (3)–(4) *is said to be strongly locally observable at* x^0 *if at* $x^0 \in \mathcal{X}$ *there is a neighborhood* $D \subset \mathcal{X}$ *such that for any* $\bar{x} \in D$, $h(F(\bar{x})) = h(F(x^0))$ *for* $k = 0, 1, \ldots, n-1$ *implies* $\bar{x} = x^0$ [19]. *System* (3)–(4) *is locally zero-state detectable in a neighborhood* D *of* $x = 0$ *if for all* $x_k \in D$, $u \equiv 0$, $y = h(\phi(k, 0, x^0, 0)) \equiv 0$ *implies that* $\lim_{k\to\infty} \phi(k, 0, x^0, 0) = 0$, $k \in \mathbb{Z}^+$.

Note that according to the notation used $x_k = \phi(k, 0, x^0, 0) = F_{k-1} \circ \cdots \circ F_0$, where $F_i = F(x_i, 0), i \in \mathbb{Z}^+$.

Theorem 4.1 ([19]) *Consider the map* $\mathcal{O} : \mathcal{X}^n \mapsto (\mathbb{R}^q)^n$ *defined by*

$$\mathcal{O}^{n-1}(x^0) = \begin{bmatrix} h(x^0) \\ h \circ F_0 \\ h \circ F_1 \circ F_0 \\ \vdots \\ h \circ F_{n-2} \circ \ldots \circ F_0 \end{bmatrix} = \begin{bmatrix} h(x^0) \\ h \circ \phi(1, 0, x^0, 0) \\ h \circ \phi(2, 0, x^0, 0) \\ \vdots \\ h \circ \phi(n-2, 0, x^0, 0) \end{bmatrix}. \tag{36}$$

If the system (3)–(4), with $u_k = 0$*, is such that it satisfies*

$$\operatorname{rank} \left[\frac{\partial}{\partial x^0} \mathcal{O}^{n-1}(x^0) \right] \Big|_{x^0 = x(0)} = n, \quad x^0 \in \mathcal{X} \tag{37}$$

then the system is strongly locally observable at $x^0 \in \mathcal{X}$.

Theorem 4.2 *If the system* (3)–(4) *satisfies* (37)*, then the system is locally zero-state observable at* x^0.

Proof The output nulling submanifold $\mathcal{N} \subset \mathcal{X}$ (see [21]) associated to the output map (4) is defined by $\mathcal{N} = \{x | h(x) = 0,\ x \in \mathcal{X}\}$. If the system (3)–(4) is such that $u_k = 0$ and $y_k = 0$, then any state trajectory evolves on $\mathcal{N}$. Any state in $\mathcal{N}$ is unobservable, since any $x \in \mathcal{N}$ (with neighborhood D) is indistinguishable from another $\bar{x} \in D \subset \mathcal{N}$. If the rank condition (37) is satisfied, necessarily $\mathcal{N} = \{0\}$, implying $x^0 = 0$. □

The previous conclusion can also be deduced from the discussion in [19]. The property of zero-state observability is important in order to assert positive definiteness of the observability function, as can be seen in the following result.

Theorem 4.3 *Assume that* (3) *with* $F(\cdot, 0)$ *is asymptotically stable on a neighborhood* D *of* $x = 0$. *If the system is zero-state observable and* L_o *exists and is smooth on* $\mathcal{X}$, *then* $L_o(x^0) > 0$, $\forall\, x^0 \in \mathcal{X}$, $x^0 \neq 0$.

Proof Recall Eq. (29), then, if $x^0 \neq 0$, zero state observability implies that for some $\bar{K} \in \mathbb{Z}^+ \backslash \{0\}$ we have $h(\phi(\bar{k}, 0, x^0, 0)) \neq 0$ for some $0 \leq \bar{k} < \bar{K}$. Therefore, if $x^0 \neq 0$, $L_o(x^0) > 0$. □

Theorem 4.4 (Existence of L_o) *If there exists a convergent* $\sum_{k=0}^{\infty} M_k$, $M_k \in \mathbb{R}$, *such that* $\|h(x_k)\|_2^2 \leq M_k$ *for all* $x_k \in D$, *then* L_o *exists as given by (33) and is a smooth solution of (32) for all* $x_0 \in D$.

Proof By Lemma 4.1, Eq. (33) is a solution of (32). Existence of the limit (33) for all $x^0 \in D$ is necessary and sufficient for existence of L_o. Since $(\mathbb{R}^n, \|\cdot\|_2)$ is a complete normed space, by Weierstrass' M test, the series of functions (34) converges uniformly and absolutely. □

With the tools developed until now, the following result can be proved and serves to establish the connection with the dissipativity theory concepts presented in the previous section.

Proposition 4.1 *Assume that the observability function* L_o *exists and is positive definite. Then* L_o *as defined in Eq. (29) is a Lyapunov function for system (3) and furthermore such system is dissipative with storage function* L_o *and supply rate* $\frac{1}{2}h^T(x_k)h(x_k)$.

Proof The proof uses similar proving techniques to those in [6] and [12] . In order to show that the difference $L_o(x_{k+1}) - L_o(x_k)$ is negative semi-definite (and thus a Lyapunov function [11]), express $L_o(x_k)$ for an arbitrary state x_k as,

$$L_o(x_k) = \frac{1}{2}h^T(x_k)h(x_k) + \frac{1}{2}\sum_{i=k}^{\infty} h^T[F(x_i, 0)]h[F(x_i, 0)],$$

doing the same for x_{k+1}, and taking the difference then

$$L_o(x_{k+1}) - L_o(x_k) = -\frac{1}{2}h^T(x_k)h(x_k), \tag{38}$$

for $k \in \mathbb{Z}^+$, which is negative semidefinite. As can be seen, the discrete-time dissipation inequality (15) is preserved and since by assumption Theorem 4.4 is satisfied, there exists M_i such that L_o is bounded, then by Theorem 3.1, L_o is a storage function with supply rate $\frac{1}{2}h^T(x_k)h(x_k)$. □

Remark 4.2 Following the terminology used in [23], Eq. (38) can be called the discrete-time Lyapunov-like equation.

To conclude this subsection, when the assumptions of Proposition 4.1 are compared with those of Theorem 3.1, the asymptotic stability and zero-state observability of system (3)–(4) seem to be natural additional requirements due to the assumption of $u_k = 0$ from the definition of L_o. The following subsection deals with the controllability function for (3)–(4).

4.2 Controllability function

Before determining some properties of the controllability function (28) of (3), it is useful to transform the definition of L_c into a more adequate representation with the help of Definition 2.2.

Remark 4.3 Consider the system (10). Then the definition of L_c from Eq. (28), may be expressed as

$$L_c(w^0) = \inf_{\substack{v\in\ell_2(0,\infty),\\ w(\infty)=0,\ w^0=w_0}} \frac{1}{2}\sum_{k=0}^{\infty} v_k^T v_k, \tag{39}$$

for w and v defined in (10).

Lemma 4.2 *Assume the existence of the optimal sequence $v^\star = \{v_i^\star | i = 0, 1, \ldots\}$ such that it satisfies* (39) *and consider the following recursive equation*

$$\mathcal{L}_c(w_{i+1}) = \mathcal{L}_c(w_i) + \frac{1}{2} v_i^{\star T} v_i^\star, \tag{40}$$

for $i = 0, 1, 2, \ldots$ and initial condition $\mathcal{L}_c(w^0) = 0$. Then $L_c(w^0)$ can be found from the solution of (40) *as follows*

$$L_c(w^0) = \lim_{i\to\infty} \mathcal{L}_c(w_i). \tag{41}$$

Proof Express (39) as,

$$L_c(w^0) = \frac{1}{2} v_0^{\star T} v_0^\star + \sum_{i=0}^{\infty} v_{i+1}^{\star T} v_{i+1}^\star, \tag{42}$$

which may be written as a recurrence equation with the initial condition $\mathcal{L}_c(w^0) = \frac{1}{2} v_0^{\star T} v_0^\star = 0$ as consequence of Remark 3.2. By solving iteratively (40), $L_c(w^0)$ can be found as i tends to infinity. □

As can be seen, Remark 4.3 and Lemma 4.2 are particular cases of Remark 3.1 and Lemma 3.2.

4.2.1 Properties of L_c

Proposition 4.2 *Assume that the system* (3) *is asymptotically stable on D that there exists a solution $v^\star$ to* (39) *and that the limit* (41) *exists. Then $L_c(w^0) > 0$ for $w^0 \in D$, $w^0 \neq 0$, if and only if the system*

$$w_{k+1} = F^{-1}(w_k, v^\star_{k+1}), \quad k \in \mathbb{Z}^+, \tag{43}$$

is asymptotically stable on D.

Proof Assume that there exists $w^0 \in D$, $w^0 \neq 0$ such that $L_c(w^0) = 0$. Since in Eq. (42) this is only possible if all $v^\star_{k+1} = 0$, for $k = 0, \ldots, \infty$, the system (43) is equivalent to the unforced system $w_{k+1} = F^{-1}(w_k, 0)$, for $k \in \mathbb{Z}^+$, but this system cannot be stable, since this would imply that the associated system $w_\kappa = F(w_{\kappa+1}, 0)$, for $\kappa \in \mathbb{Z}^-$ is unstable, which contradicts the asymptotic stability of F. □

Proposition 4.3 *A necessary existence condition of $L_c(w_k)$ in Eq.* (40) *is that $v^\star_k$ is the solution of the following two-point boundary value problem*

$$\lambda_k = \left[\frac{\partial}{\partial w_k} F^{-1}(w_k, v_{k+1})\right]^T \lambda_{k+1}, \tag{44}$$

$$v_{k+1} = -\left[\frac{\partial}{\partial v_{k+1}} F^{-1}(w_k, v_{k+1})\right]^T \lambda_{k+1}, \tag{45}$$

subject to the boundary conditions $w(\infty) = 0$ and $w^0 = w(0)$.

Proof In order to find $L_c(w_k)$ given by Eq. (39), applying standard tools of the discrete optimal control theory (see for instance [1,13]) results in the following Hamiltonian,

$$H_k = \frac{1}{2} v^T_{k+1} v_{k+1} + \lambda^T_{k+1} F^{-1}(w_k, v_{k+1}), \tag{46}$$

resulting in

$$\frac{\partial H_k}{\partial w_k} = \lambda^T_{k+1} \frac{\partial}{\partial w_k} F^{-1}(w_k, v_{k+1}) = \lambda^T_k,$$
$$\frac{\partial H_k}{\partial v_{k+1}} = v^T_{k+1} + \lambda^T_{k+1} \frac{\partial}{\partial v_{k+1}} F^{-1}(w_k, v_{k+1}) = 0,$$

from which Eqs. (44) and (45) follow. □

As can be observed from Eq. (45), the input v_{k+1} may appear implicitly. Therefore, the analytical solution of this problem may be difficult to find in the nonlinear case. In the Appendix, this optimal control problem is presented in a sequential fashion in order to show the structure of this problem. In contrast for linear systems Eqs. (44) and (45) can be solved explicitely, as can be seen in the following example.

Example 4.2 An alternative proof of Corollary 4.1 to find L_c is provided here. Assume the existence of A^{-1} and consider the system from Def. 2.2 associated to Eq. (8) and provided in Eq. (13) whose general solution can be expressed as

$$w_k = A^{-k}w^0 - \sum_{i=0}^{k-1}(A^{-1})^{k-i}Bv_{i+1}. \tag{47}$$

Using (44) and (45), results in

$$\lambda_k = A^{-T}\lambda_{k+1}, \tag{48}$$
$$v_{k+1} = B^T A^{-T}\lambda_{k+1}. \tag{49}$$

Substitution of (49) in (13) yields,

$$w_{k+1} = A^{-1}w_k - A^{-1}BB^T A^{-T}\lambda_{k+1}. \tag{50}$$

Solving Eq. (48) explicitly in *backward time*, results in

$$\lambda_k = (A^{-T})^{N_p-k}\lambda_{N_p}. \tag{51}$$

Then the solution of (50) with input λ_{k+1} given by (51) is

$$w_k = A^{-k}w_0 - \sum_{i=0}^{k-1}A^{i-k}BB^T(A^T)^{i-N_f}\lambda_{N_p}. \tag{52}$$

For $w_{N_p} = 0$, Eq. (52) implies that, $w^0 = P(A^T)^{-N_p}\lambda_{N_p}$ where $P = \sum_{i=0}^{N_p-1} A^iBB^T(A^T)^i$, which can be expressed as $\lambda_{N_p} = (A^T)^{N_p}P^{-1}x^0$, which in Eq. (51) for λ_{k+1} and this result in Eq. (49), yields $v^\star_{k+1} = B^T(A^T)^kP^{-1}w^0$ which after substitution in Eq. (40) results in Eq. (30).

Theorem 4.5 (Existence of L_c) *Assume that $v^\star$ satisfies Eq.* (39) *with $L_c(w^0)$ smooth for all $x^0 \in D$ and such that Eq.* (43) *is asymptotically stable. Let $\|v^\star_i\|_2^2 \leq M_i$, $M_i \in \mathbb{R}$ such that $\sum_{i=0}^{\infty} M_i$ converges uniformly and absolutely. Then $L_c(w^0)$ exists as given by* (41) *and is a smooth solution of* (40) *for all $w^0 \in D$.*

Proof By Remark 4.3 existence of $L_c(x^0)$ is equivalent to existence of $L_c(w^0)$. By Lemma 4.2, Eq. (41) is a solution of (40). $L_c(w^0)$ exists if the series of functions (41) converges. Since $(\mathbb{R}^n, \|\cdot\|_2)$ is a complete normed space, by Weierstrass' M-Theorem, the series (41) converges uniformly and absolutely. □

The latter results serve to prove the following proposition, which establishes the connection with the concepts of dissipativity theory of the previous section.

Proposition 4.4 *Assume that the conditions of Theorem* 4.5 *are satisfied, then the controllability function* $L_c(w^0)$ *as defined in Eq.* (39) *is a Lyapunov function for system* (10). *Furthermore, the system* (10) *is dissipative and* $L_c(w_k)$ *is also a storage function, with supply rate* $\frac{1}{2}v_k^{\star T}v_k^{\star}$.

Proof That $L_c(w_k)$ is a Lyapunov function for (10), can be shown noticing its nonnegative definitness from Eq. (39). By Proposition 4.2 $L_c(w^0) > 0$ for $w^0 \in D$. In order to show that the difference $L_c(w_{k+1}) - L_c(w_k)$ is negative semidefinite, note that for an arbitrary state w_k, from (42), L_c can be expressed as

$$L_c(w_k) = \frac{1}{2}v_k^{\star T}v_k^{\star} + \frac{1}{2}\sum_{i=k}^{\infty} v_{i+1}^{\star T}v_{i+1}^{\star}, \tag{53}$$

doing the same for w_{k+1}, and taking the difference yields,

$$L_c(w_{k+1}) - L_c(w_k) = -\frac{1}{2}v_k^{\star T}v_k^{\star}, \tag{54}$$

which is negative semidefinite. Since the discrete-time dissipation inequality (15) is preserved and by Theorem 4.5 $L_c(w_k)$ is bounded, then $L_c(w_k)$ is a storage function with supply rate $\frac{1}{2}v_k^{\star T}v_k^{\star}$. □

Some comments about the latter results are pertinent. A comparison of the assumptions of Proposition 4.4 with those of Theorem 3.2 reveals that asymptotic stability of system (3) is a stronger assumption than just dissipativity. However, the same differences can be observed in the continuous-time case, see [24,27].

The existence condition of Theorem 4.5 is useful in order to provide solvability conditions for the dynamic optimization problem (26)–(27) restricted to the supply rate $r_r = u_k^T u_k$, which finally results in the optimization problem of the following section.

4.2.2 Optimization-based search of $v_k^{\star}$

Define the finite set $\{v_i | i = 0, \ldots, N_p\} \subset \{v_i | i = 0, \ldots \infty\}$ such that Eq. (39) is satisfied. Then by using a dynamic optimization approach [1], the optimization problem takes the form

$$\min_{\{v_i | i=1,\ldots,N_p\}} \frac{1}{2}\sum_{i=0}^{N_p} v_i^T v_i,$$

with equality constraints $w_{N_p} = 0$ and $w^0 = w_k$, which is expressed in the form of the Mayer problem (see e.g., [1]).

Assume that the conditions of Theorem 4.5 are satisfied. Let $\epsilon, N_p \in \mathbb{Z}^+$ be such that $\|w_{N_p}\| \leq \epsilon$ for ϵ small enough. Assume that N_p is known and the (closed) set of admissible inputs $\{v \mid v \in \mathcal{V}\}$ is convex with nonempty interior. Then $L_c(w_k)$ in Eq. (40) can be determined with the solution of the following optimization problem

$$\min_{\{v_i | i=1,\ldots,N_p\}} \mathcal{L}_c(w_{N_p}), \tag{55}$$

with equality constraints

$$\begin{cases} w_{i+1} = F^{-1}(w_i, v_{i+1}), \\ \mathcal{L}_c(w_{i+1}) = \mathcal{L}_c(w_i) + \frac{1}{2} v_i^T v_i, \\ w_{N_p} = 0, \\ w^0 = w_k, \end{cases} \tag{56}$$

with initial inputs $\{v^0 \mid v_j^0 \in \mathcal{V},\ j = 1, \ldots, N_p\}$ and with $\mathcal{L}_c(w^0) = 0$, determining $v_i^\star$.

A drawback of this approach can be pointed out. Though for an asymptotically stable system N_p can be approximated to be finite, introducing some error in the result, the *best value* of N_p is unknown prior to the nonlinear optimization process.

4.3 Example: the energy functions of a nonlinear system

In order to illustrate some advantages and limitations of the approach previously presented, we find the energy functions associated to a simple model of a series interconnected motor (also known as universal motor).

When departing from a discrete-time system, invertibility of the discrete-time map $F(\cdot, u_k)$ is an assumption required for reversed-time evolution. A simple discrete-time example can be found in [14].

In this example we depart from a continuous-time system and therefore the map $F(\cdot, u_k)$ is a diffeomorphism [2]. In consequence, the controllability function can be found using Remark 2.1 circumventing the need of explicitely inverting such map. Consider the universal motor depicted in Fig. 2 with specifications defined in Table 1.

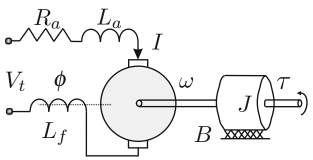


Fig. 2 Universal motor

Table 1 Specifications of the universal motor

Variable	Value	Units
Resistance R	2.5	Ω
Inductance (field + armature) ($L = L_f + L_a$)	0.08	H
Rotational inertia (J)	10	$\mathrm{Kg\,m^2\,s/rad^2}$
Angular position(θ)	$\theta \in \mathbb{R}$	rad
Angular velocity (ω)	$(I, \omega) \in \mathcal{Y}$	rad/s
Current (I)	$(I, \omega) \in \mathcal{Y}$	A
Voltage at terminals (V_t)	$V_t \in \mathcal{U}$	V
Rotational damping (B)	0.50	N m s/rad
Constant (torque) (K_T)	0.42	N m / Wb A
Constant (MMF) (K_m)	0.42	V s / Wb rad
Constant (field) (K_f)	0.53	Wb/A

The dynamic behavior of this system may be described by

$$\begin{cases} L\frac{\mathrm{d}I}{\mathrm{d}t} = V_t - RI - \zeta\omega I, \\ J\frac{\mathrm{d}\omega}{\mathrm{d}t} = \zeta I^2 - B\omega, \end{cases} \qquad t \in \mathbb{R} \tag{57}$$

with state defined by $x = (I, \omega)$, $x \in \mathcal{X}$ with input V_t and outputs I and ω. Considering that $K_m = K_t$, define $\zeta = K_t K_f$. Although this system is *locally accessible* it is not controllable, since $\dot{\theta} = \omega \geq 0 \ \forall \ x \in \mathcal{X}$. The corresponding controllability function has a region where $L_c = 0$ in one half of the state space with a nonsmooth transition to the other half. While such difficulties are essentially unimportant for the optimization approach, such energy function cannot be approximated adequately by analytic functions limiting the computer representation and manipulation of such functions. For instance for a polynomial fit the approximation of such energy functions may be unsuccessful. For convenience define $a = \frac{-R}{L}$, $b = \frac{-\zeta}{L}$, $c = \frac{\zeta}{J}$, $d = \frac{-B}{J}$ and $e = \frac{1}{L}$. By the Taylor–Lie series, a discrete-time approximation of system (57) is given by

$$\begin{aligned} \begin{bmatrix} I \\ \omega \end{bmatrix}_{k+1} &= \begin{bmatrix} I \\ \omega \end{bmatrix}_k + \mathcal{T} \begin{bmatrix} aI + bI\omega + eV_t \\ cI^2 + d\omega \end{bmatrix}_k \\ &\quad + \frac{\mathcal{T}^2}{2} \begin{bmatrix} a + b\omega & bI \\ 2cI & d \end{bmatrix} \begin{bmatrix} aI + bI\omega + eV_t \\ cI^2 + d\omega \end{bmatrix}_k + \cdots \end{aligned} \tag{58}$$

with outputs I_k and w_k. Following the sequence of steps described in Definition 2.1, the associated system of (57) is in this case,

$$\begin{cases} L\frac{\mathrm{d}I}{\mathrm{d}t} = RI + \zeta\omega I - V_t, \\ J\frac{\mathrm{d}\omega}{\mathrm{d}t} = B\omega - \zeta I^2, \end{cases} \qquad t \in \mathbb{R}^+ \tag{59}$$

and due to the commutativity in the diagram of Fig. 1 the associated discrete-time system for a state $w_k = (\bar{I}_k, \bar{\omega}_k)$ is given by

$$\begin{bmatrix} \bar{I} \\ \bar{\omega} \end{bmatrix}_{\kappa+1} = \begin{bmatrix} \bar{I} \\ \bar{\omega} \end{bmatrix}_k + \mathcal{T} \begin{bmatrix} -a\bar{I} - b\bar{I}\bar{\omega} - ev \\ -c\bar{I}^2 - d\bar{\omega} \end{bmatrix}_k + \frac{\mathcal{T}^2}{2} \begin{bmatrix} -a - b\bar{\omega} & -b\bar{I} \\ -2c\bar{I} & -d \end{bmatrix} \begin{bmatrix} -a\bar{I} - b\bar{I}\bar{\omega} - ev \\ -c\bar{I}^2 - d\bar{\omega} \end{bmatrix}_k + \cdots \tag{60}$$

where the input is $v_k = \bar{V}_t(k+1)$ and the outputs are $\bar{I}_k$ and $\bar{w}_k$. In order to determine L_c and L_o the values of the parameters shown in Table 1 were used. All the routines and graphics of this example were performed using Matlab.

Observability function Consider the iterative solution of Eq. (32), as $i \to \infty$ for each initial state x_0 within the desired region to plot. The resulting observability function is given in Fig. 3a. It can be seen from the graphic in Fig. 3a that L_o is more influenced by I than by ω.

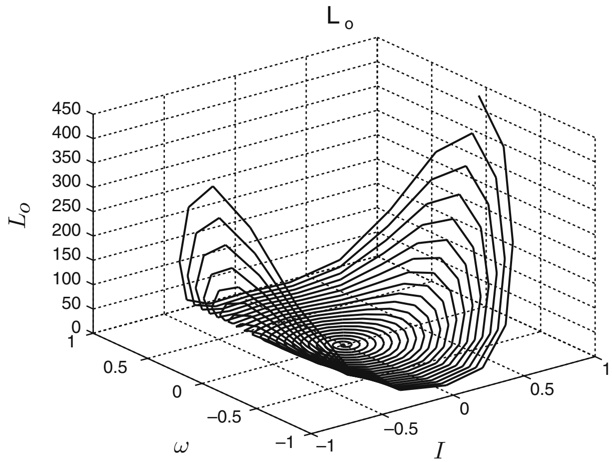
Controllability function By using the optimization approach of Sect. 4.2.2 and defining a finite set $\{v_i | i = 1 \ldots N_p\}$, the optimization problem stated in Eq. (55),(56) can be solved for each selected state w of (60) and thus the results can be plotted resulting in Fig. 3b. The region where $L_c = 0$ in one half of the state space which corresponds to the unreachable part where $\dot{\theta} = \omega \leq 0$ for $x \in X$ can be seen from the figure. Also it can be discerned that several points did not converge to the region where $L_c = 0$. The Optimization Toolbox of Matlab was used to find $v^\star$ using an spiraled grid with the purpose of simplifying the optimization task until an adequate radius is obtained.

5 Conclusions

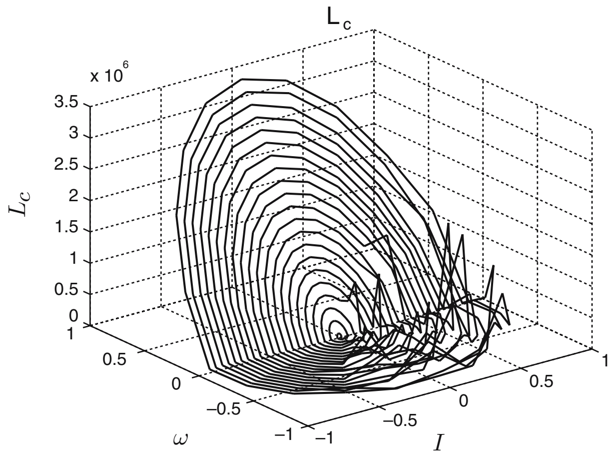
Dissipativity theory is a general framework useful to establish input-state-output relationships and with this, to pose several storage functions used in nonlinear balancing. Using properly defined dynamic optimization problems, along with adequate nonlinear discretization algorithms – including those based on Taylor–Lie series [9] or numerical integration algorithms – it is possible to provide a framework to find approximations of such storage functions.

In particular, the discrete-time versions of the controllability and observability energy functions were discussed. Instead of looking for the solution of a Hamilton–Jacobi–Isaacs and a Lyapunov-like partial differential equations as in the continuous-time case, an optimization approach and an iterative algorithm were proposed to find L_c and L_o, respectively. This approach was exemplified with linear and nonlinear discrete-time systems. In particular using this approach on the discrete-time equivalent model of a universal motor the approximated energy functions were found.

Topics of ongoing research are the further development of the nonlinear approach of dissipative balancing. Of particular interest for us is the case of port-Hamiltonian systems since the interaction of ports with the state can be nat-



(a) *Observability function* L_o



(b) *Controllability function* L_c

Fig. 3 Energy functions for the universal motor

urally posed in this framework, establishing relations with the physical energy stored in the system [15].

Appendix

Proofs of Theorems 3.1 and 3.2

Proof (Theorem 3.1). Suppose that $S_a(x^0, r_a)$ is finite. That $S_a(x^0, r_a) \geq 0$ can be verified by taking $N_f = 0$ in (17). Consider the value of $S_a(x^0, r_a)$ at two points x_{k+m} and x_k located at the trajectory defined by the optimal sequence of inputs $\{u_i^\star | i = 0, 1, \ldots, N_f - 1\}$ that satisfy (18), then the difference can be expressed as $S_a(x_{k+m}, r_a) - S_a(x_k, r_a) = \sum_{i=k}^{k+m-1} r_a(y_i, u_i^\star)$. In any other suboptimal trajectory we have $\sum_{i=k}^{k+m-1} r_a(y_i, u_i) \geq \sum_{i=k}^{k+m-1} r_a(y_i, u_i^\star)$, resulting in $S_a(x_{k+m}, r_a) - S_a(x_k, r_a) \leq \sum_{i=k}^{k+m-1} r_a(y_i, u_i)$ satisfying (16). Assume now that (3)–(4) is dissipative. Then there exist some $S(x_k, r) \geq 0$ that for any u_k satisfies (16). Since for $x_k = x^0$, we have $S(x^0, r) + \sum_{i=0}^{N_f} r(y_i, u_i) \geq S(x_{N_f}, r) \geq 0$, which after comparison with (17) it may be seen that $S(x^0, r) \geq S_a(x^0, r_a) \geq 0$ and therefore $S_a(x^0, r_a)$ must be finite and bounded from above by any other storage function. □

Proof (Theorem 3.2) Consider the value of $S_r(x^0, r_r)$ at two points x_k and x_{k+m} located at the trajectory defined by the optimal sequence of inputs $\{u_i^\star | i = -N_p, \ldots, -1, 0\}$ that satisfy (22) and departs from $x_{-N_p} = x^*_{-N_p}$ towards x^0. The difference $S_r(x_{k+m}, r_r) - S_r(x_k, r_r)$ is given by $S_r(x_{k+m}, r_r) - S_r(x_k, r_r) = \sum_{i=k+1}^{k+m} r_r(y_i, u_i^\star)$, while for any other suboptimal trajectory $\sum_{i=k+1}^{k+m} r_r(y_i, u_i) \geq \sum_{i=k+1}^{k+m} r_r(y_i, u_i^\star)$, and therefore $S_r(x_{k+m}, r_r) - S_r(x_k, r_r) \leq \sum_{i=k+1}^{k+m} r_r(y_i, u_i)$ satisfying (16). At the point x^* the following relation holds $S_a(x^*, r_a) = \sup_x -S_r(x, r_r)$, [27]. Since reachability implies the possibility of steering x^* to x in finite time, in order to have $S_a(x^*, r_a)$ finite, there must exist a bound M for $S_r(x, r_r)$ such that $-\infty < M \leq S_r(x, r_r)$, concluding the proof. □

About the structure of $v_k^\star$

In this appended subsection, some difficulties of the optimal control discussed in Proposition 4.3 are reviewed. In order to study the structure of $v_k^\star$ in (44)–(45), the corresponding boundary value problem is addressed. Some definitions are useful to simplify notation. A successive composition of functions is denoted as, $F_{[m,n]} \overset{\text{def}}{=} F_m \circ F_{m+1} \circ \cdots \circ F_n$, with $F_{[n,n]} \overset{\text{def}}{=} F_n$. With a slight abuse of notation, $F_i^{-1} \overset{\text{def}}{=} F^{-1}(x_i, v_{i+1})$.

Define the following maps,

$$\Psi_k = \frac{\partial}{\partial w_k} F^{-1}(w_k, v_{k+1}), \tag{61}$$

$$\Upsilon_k = -\frac{\partial}{\partial v_{k+1}} F^{-1}(w_k, v_{k+1}), \tag{62}$$

and denote by $\Psi_{[m,n]} \overset{\text{def}}{=} \Psi_m \Psi_{m+1} \cdots \Psi_n$ the successive application of a step variant linear map Ψ_k for a discrete interval $k \in [m, n]$. Then the solution of (44), given an initial λ_{N_p}, with $0 \leq k \leq N_p$ can be expressed as, $\lambda_k = \Psi^T_{[k,N_p-1]} \lambda_{N_p}$, and in consequence the possibly implicit input v_{k+1} can be obtained from the following expression,

$$v_{k+1} = \Upsilon_k^T \Psi_{[k+1,N_p-1]}^T \lambda_{N_p}. \tag{63}$$

Consider the following composition operations for the map $F_{[i,N_p]} \stackrel{\text{def}}{=} F_{i+1} \circ F_{i+2} \circ \cdots \circ F_{N_p}$ and for the inverse map $F_{[i,0]}^{-1} \stackrel{\text{def}}{=} F_i^{-1} \circ F_{i-1}^{-1} \circ \cdots \circ F_0^{-1}$. Then both Eq. (10) and the system $w_\kappa = F(w_{\kappa+1}, v_{\kappa+1})$, $\kappa \in \mathbb{Z}^-$, which evolves in backward-time, can be expressed in terms of equation (63) as, $w_{\kappa+1} = F^{-1}(w_\kappa, \Upsilon_\kappa^T \Psi_{[\kappa+1,N_p-1]}^T \lambda_{N_p})$, and $w_\kappa = F(w_{\kappa+1}, \Upsilon_\kappa^T \Psi_{[\kappa+1,N_p-1]}^T \lambda_{N_p})$. At the boundary for $\kappa = 0$, $w(0) = w_0$, $w_0 = F(w_1, \Upsilon_0^T \Psi_{[1,N_p-1]}^T \lambda_{N_p}) = F_{[0,N_p]}$, and for $\kappa = N_p$, $w_{N_p} = 0$,

$$0 = F^{-1}(w_{N_p-1}, \Upsilon_{N_p-1}^T \Phi_{N_p-1}^T \lambda_{N_p}) = F_{[N_p,0]}^{-1}, \tag{64}$$

and its inverse map is $w_{N_p-1} = F(0, \Upsilon_{N_p-1}^T \Phi_{N_p-1}^T \lambda_{N_p}) = F_{[N_p,N_p]}$. In this last equation, we have a nonlinear relation between w_{N_p-1} and λ_{N_p}. Notice also that $\Upsilon_{N_p-1} = \Upsilon(w_{N_p-1}, v_{N_p})$, with v_{N_p} inserted possibly in implicit form. In the linear case this never occurs and it is solvable, see Example 4.2. In the general case this problem is difficult to solve in closed form. However, as shown throughout the paper, dynamic optimization algorithms can be used in order to solve it.

Acknowledgement This work is partially supported by Instituto Mexicano del Petróleo (IMP), Eje Central Lázaro Cárdenas no. 152, Col. San Bartolo Atepehuacan C.P. 07730, Mexico City, Mexico, Apartado Postal 14-805. http://www.imp.mx

References

1. Bryson AE Jr (1999) Dynamic optimization. Addison-Wesley Longman, USA
2. Fliess M (1992) Reversible linear and nonlinear discrete-time dynamics IEEE Trans Aut Cont 37(8): 1144–1153
3. Fujimoto K, Scherpen JMA (2005) Nonlinear input-normal realizations based on the differential eigenstructure of Hankel operators IEEE Trans Aut Cont AC-40(1): 2–19
4. Glover K (1984) All optimal Hankel norm approximations of linear multivariable systems and their L_∞ error bounds Int J Control 39: 1115–1193
5. Gray WS, Mesko JP (1999) Observability functions for linear and nonlinear systems. Syst Cont Lett 38: 99–113
6. Guillard H, Monaco S, Normand-Cyrot D (1996) On H_∞ Control of discrete-time nonlinear systems Int. J. Robust Nonlinear Cont 6: 633–643
7. Grizzle JW (1993) A linear algebraic framework for the analysis of discrete-time nonlinear systems. SIAM J Cont Optim 31(4): 1026–1044
8. Jakubczyk B, Sontag ED (1990) Controllability of nonlinear discrete-time systems: a Lie-algebraic approach. SIAM J Cont Optimiz 28(1): 1–33
9. Kazantzis N, Kravaris C (1997) System-theoretic properties of sample-data representations of nonlinear systems obtained via Taylor-Lie series Int J Control 67(6): 997–1020
10. Kotta U (1995) inversion method in the discrete-time nonlinear control systems synthesis problems. Lect N Cont Inf Sci Springer, Berlin Heidelberg New York p 205
11. LaSalle JP (1976) The stability of dynamical systems. SIAM, Philadelphia
12. Lin W, Byrnes CI (1996) $\mathcal{H}_\infty$- control of discrete-time nonlinear systems IEEE Trans Aut Cont 41(4): 494–510
13. Lewis FL, Syrmos VL (1995) Optimal Control, 2nd edn. John Wiley & Sons, USA

14. Lopezlena R, Scherpen JMA (2002) Energy functions and balancing for discrete -time nonlinear systems In: Proc. IFAC World Congress, Barcelona, Spain
15. Lopezlena R, Scherpen JMA, Fujimoto K (2003) Energy-storage balanced reduction of port-Hamiltonian systems In 2*nd* IFAC Workshop on Lagr and Ham Meth for Nonl Cont pp. 79–84, Seville, Spain
16. Monaco S, Normand-Cyrot D (1986) Nonlinear systems in discrete time. In: Fliess M, Hazewinkel (eds) Algebraic and geometric methods in nonlinear control theory. Reidel, The Netherlands pp. 411–430
17. Monaco S, Normand-Cyrot D (1997) A Unified representation for nonlinear discrete-time and sampled dynamics J Math Syst Est Cont 7(4): 477–503
18. Moore BC (1981) Principal components analysis in linear systems: controllability, observability, and model reduction IEEE Trans Aut Cont AC-26(1): 17–32
19. Nijmeijer H (1982) Observability of autonomous discrete time non-linear systems: a geometric approach Int J Control 36(5): 867–874
20. Newman AJ, Krishnaprasad PS (1998) Computation for nonlinear balancing In: Proceedings. 37*th* IEEE Conference on Dec & Cont., Tampa Fl USA
21. Nijmeijer H, van der Schaft AJ (1990) Nonlinear dynamical control systems. Springer, Berlin Heidelberg New York
22. Pernebo L, Silverman LM (1982) Model reduction via balanced state space representations IEEE Trans Aut. Cont. 27: 382–387
23. Scherpen JMA (1993) Balancing for nonlinear systems Syst Contr Lett 21: 143–153
24. Scherpen JMA (1994) Balancing for Nonlinear Systems. Ph.D. thesis, University of Twente
25. Scherpen JMA (1996) $\mathcal{H}_\infty$-balancing for nonlinear systems. Int J Rob Nonl Contr 6: 645–668
26. Scherpen JMA, van der Schaft AJ (1994) Normalized coprime factorizations and balancing for unstable nonlinear systems Int J Contr 60(6): 1193–1222
27. van der Schaft AJ (2000) L_2-gain and passivity techniques in nonlinear control. Springer, Berlin Heidelberg New York
28. Verriest EI, Gray WS (2001) Discrete-time nonlinear balancing In: Proceedings of 5th IFAC-Nonlinear Cont Syst Des Symp St. Petersburg, Russia pp. 515–520
29. Weiland S (1991) Theory of approximation and disturbance attenuation for linear systems. Ph.D. dissertation, Rijksuniversiteit Groningen
30. Weiland S (1994) A behavioral approach to balanced representations of dynamical systems. Lin Alg Appl 205–206: 1227–1252
31. Willems JC (1972) Dissipative dynamical systems, Part 1: general theory Arch Rat Mech Anal 45: 321–351